Comparison of Higher-Order Ambisonics, Vector- and Distance-Based Amplitude Panning using a hearing device beamformer

Simon, Laurent S R ; Wuethrich, Hannes ; Dillier, Norbert

Comparison of Higher-Order Ambisonics, Vector- and Distance-Based Amplitude Panning using a hearing device beamformer

Laurent S. R. Simon[1], Hannes Wuethrich[2], and Norbert Dillier[3]

[1] *ORL Klinik, University Hospital, Zürich, Switzerland, Email: laurent.simon@usz.ch*

[2] *Sonova AG, Stäfa, Switzerland, Email: hannes.wuethrich@sonova.com*

[3] *ORL Klinik, University Hospital, Zürich, Switzerland, Email: norbert.dillier@usz.ch*

Abstract

The main advantages and drawbacks of Higher Order Ambisonics (HOA) versus Vector-Based Amplitude Panning (VBAP) for 3D audio sound reproduction are well documented. However, few studies focused on the effect of 3D audio reproduction artifacts and errors on the behaviour of hearing devices. This study compares the effect of HOA, VBAP, and Distance-Based Amplitude Panning (DBAP) sound reproduction on the 3D directivity index of a hearing device beamformer at the sweet spot and for two off-centered positions in a non-anechoic room using 29 loudspeakers. Preliminary results at the sweet spot show little difference between the directivity patterns obtained with HOA, VBAP, DBAP, and real sources. At off-centered positions, the difference between HOA and real sources increases.

Introduction

According to [1], the most important audio quality of hearing aids is to improve the clarity of sound. Clarity of sound is generally evaluated through speech intelligibility tests [2, 3], the main type of test for evaluation of hearing aids.

However, speech intelligibility tests provide no information about spatial perception: localization, immersion, externalization, ... In addition, speech intelligibility tests are for now limited to the horizontal plane, and therefore do not cover cases where the hearing aid user is facing downwards (e.g. when eating or taking notes during class). In order to overcome this limitation, recent work has focused on evaluation of hearing devices in 3D audio environments [4, 5, 6] or in realistic environments [7, 8].

In order to simulate an auditory scene, one can find three major families of 3D audio reproduction techniques: binaural, wave field reconstruction, and stereophonic techniques.

Binaural techniques aim to reproduce at a listener's ear canals the audio signals that would have otherwise arrived at their ears if they had been in the middle of a real auditory scene [9]. A number of factors limit the good reproduction of auditory scenes via binaural techniques. It is necessary to use individual Head-Related Transfer Functions (HRTF), the filters that simulate the path from a virtual source to each of the listener's ears, and their measurement is time-consuming. Unless a headtracker is used, head movements are not taken into account, which causes front-back confusions and reduces externalization [10].

Wave field reconstruction techniques, such as Higher Order Ambisonics (HOA) or Wave Field Synthesis (WFS), aim to reproduce a physical copy of a given sound field over an extended area, using loudspeakers. It may seem the most appropriate type of 3D sound reproduction technique for hearing device evaluation: as [11] pointed out, hearing devices do not work the same way as the human brain. Before amplifying the sound that it picks up and reproduce it to its owner, a hearing aid processes it through different directional filtering, classification, amplification, compression and denoising stages [12]. In that condition, using a 3D audio reproduction system that is perceptually-based could limit the usability of the hearing aid evaluation results.

Stereophonic techniques are loudspeaker-based techniques that produce phantom sources between loudspeakers through the use of time and level differences between loudspeakers. They differ from wave field reconstruction techniques by the fact that they are based not on physical but perceptual rules, e.g. Vector-Based Amplitude Panning (VBAP).

These techniques were perceptually and physically evaluated in the past with normal hearing listeners [13, 14, 15, 16] but evaluations of these techniques in a hearing aid context are sparse.

In the context of hearing devices, binaural techniques require to use hearing aid satellites, limiting testing to hearing aid prototypes. With binaural reproduction, it is therefore impossible to use a listener's own hearing aids.

Past studies on 3D audio for hearing devices therefore focused on HOA [17, 18] at the sweet spot, i.e. at the ideal listening point of the system. In [6], Grimm simulated VBAP, HOA and nearest loudspeaker reproductions using a horizontal loudspeaker array. Measurements were performed by convolving the loudspeaker signals with Head-Related Impulse Responses (HRIR) measured in an anechoic room using hearing aid satellites. The simulated listening positions were the center of the loudspeaker array, 10cm off-centered, and 50cm off-centered. However, the off-centered positions were simulated using a correction of the gains and delays of the loudspeaker channels and an interpolation of the HRIR between the measurement positions, not taking into account the change of directivity of the loudspeakers.

However, using HOA, all loudspeakers are active all the time. The level of the loudspeakers close to the intended direction of the virtual source is higher than that of further loudspeakers, but all loudspeakers contribute to the simulation of the sound field. When the sound field is perfectly reconstructed at the position of the hearing aid, this should not be problematic. However, if the sound field reconstruction conditions are not ideal (e.g. if a head is present in the sound field, if the order of the HOA encoding / decoding is too low, if the hearing aid is not located at the sweet spot, or if the disposition of the loudspeakers is sub-optimal for the decoding), the performance of the beamformer could be decreased [11].

VBAP uses at maximum 2 loudspeakers at a time in 2D, and maximum 3 at a time in 3D. If the hearing aids are away from the sweet spot, the beamformer will therefore pick up sounds that are coming from a less erroneous direction with VBAP than with HOA. For this reason, it was expected that the performance of the hearing aid beamformer would be better with VBAP than with HOA, especially for off-centered positions and at high frequencies. VBAP's main perceptual drawback compared to HOA when using moving sources is that virtual sources tend to "jump" from one loudspeaker to the other [19].

Distance-Based Amplitude Panning (DBAP) is an alternative 3D audio reproduction method whose processing is independent of the listener position [20]. The level of each loudspeaker is inversely proportional to the distance between the virtual source and the loudspeaker. This reproduction technique is initially aimed at concert halls and large listening areas, as it does not assume a precise listener position. It was nevertheless considered for the current study, one of the experiment variables being the distance of the listener to the sweet spot, as discussed below.

Aside of [6], all past studies on 3D audio reproduction systems done in the context of hearing aids were done only at the sweet spot. Forcing the listener to keep its head centered heavily constrains its movements. Such a limitation of movements can be problematic in an experiment that aims to reproduce natural listening conditions.

Additionally, most previous studies on the use of Virtual Sound Environments (VSE) for the evaluation of hearing aids were either done in anechoic rooms or using simulations of sound reproduction systems [11, 6]. The current study focuses on 3D audio reproduction systems in a real room, similar to that used in [17]. The behaviour of a monaural, dual-microphone MVDR beamformer (Minimum Variance Distortionless Response beamformer) is compared for VBAP, HOA, DBAP, and real sources, for an artificial head located either at the center of the loudspeaker array, 10cm off-centered, or 20cm off-centered.

In the Experimental Conditions section, we describe the loudspeaker setup, the 3D reproduction algorithms, as well as the measurement conditions. The Results section describes the results of the measurements. Finally, these results are discussed in comparison to previous studies and hearing devices algorithms.

Experimental conditions

Previous studies on the use of VSE for hearing devices evaluation used anechoic rooms. Although the use of anechoic rooms is ideal, it is also more costly and less common than using a room with a low reverberation time. For practical reasons, the measurements therefore took place in a dry room designed for listening tests, with an RT30 of 117ms. It was therefore expected that the Directivity Index (DI) of the tested beamformer would be lower than in an anechoic room, both for real and virtual sources.

The loudspeaker setup consisted of 29 Meyer Sound MM4 Xp loudspeakers, shown in Figure 1:

- 8 loudspeakers every 45° azimuth at an elevation of -38°
- 12 loudspeakers every 30° azimuth at an elevation of 0°
- 8 loudspeakers every 45° azimuth at an elevation of 39°
- 1 loudspeaker at an elevation of 90°

Positions of the loudspeakers were constrained by the small dimensions of the listening room. Distance from the loudspeakers to the center of the array was 1.5m, except for the top loudspeaker, which was 0.92m away from the center.

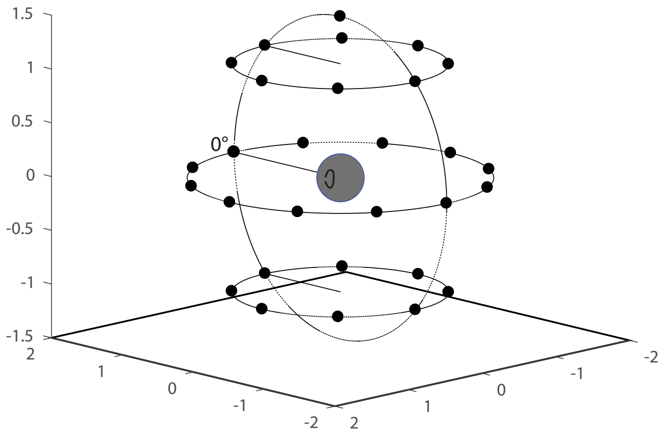

Figure 1: Loudspeaker positions used for this study, in meters. Plain straight lines indicate the 0° direction.

The levels and delays of all the loudspeakers were corrected during a calibration process, to make the sound from all loudspeakers arrive at the same time and level at the sweet spot.

3D audio encoding and decoding was performed in Spat [21] with MAX/MSP. Analysis of the measurements was performed in Matlab.

For this study, all Spat room effects were removed from the processing chain.

VBAP and DBAP decoding used Spat default parameters.

For HOA, several decoding methods and decoding types were tested, using a third order HOA decoding. Both Mode-Matching and Energy-preserving decoders [22, 23] were tested, using either basic or $max - r_E$ weighting [24].

The maximum HOA reconstruction frequency is given by equation (1)

$$f_{HOA} \leq \frac{c}{2\Pi r} M \qquad (1)$$

where c is the velocity of sound in the air, r the diameter of the listening area, and M the order of the decoding [11]. For $r = 8.5cm$, an area is defined that would contain an average size head without being allowed to move, $f_{cut} =$ 1909 Hz. Above this frequency, the reconstruction error will become larger than -14dB, which according to [11, 25, 26], should be sufficient for most hearing aids applications.

Impulse response measurements were done using KEMAR and a dual-microphone Behind-The-Ear hearing aid satellite, i.e. a hearing aid without any processing. The hearing aid satellite was fitted to KEMAR's left ear. The measurements used 2s sweeps ranging from 100Hz to 20kHz. Head-Related Transfer Functions (HRTFs) were measured with the Meyer Sound loudspeakers in an anechoic room (real sources only), to serve as a reference, and in the measurement room (real and virtual sources).

The beamformer filter was estimated using the (0° elevation, 0° azimuth) anechoic impulse response as the desired source and the (0° elevation, 180° azimuth) anechoic impulse response as the noise source.

HRTFs were measured using the same hearing aid satellites and KEMAR manikin for each of the loudspeakers and for the virtual sources produced by the DBAP, the HOA, and the VBAP systems. For each 3D audio reproduction system, virtual sources were produced every 10° in elevation between -40° and +80°, and every 10° in azimuts between 0° and +350°, resulting in 29 real source positions and 468 virtual source positions.

Results

The DI was measured using all the available measurement points for each system, meaning that the DI measured for real sources used only 32 positions whereas the DI measured for the DBAP, HOA, and VBAP used 468 positions. Using the same positions for the DI estimation would have biased the measurement, as for VBAP, for example, the virtual sources are identical to the real sources at the positions of the loudspeakers, resulting in a beamformer behaving better at the positions of the loudspeakers than in between. DI's were measured on a cone, integrating the impulse response measurements over an angle of 60° centered on the 0° azimuth and 0° elevation. At each frequency, the gain of the beamformer at 0° azimuth and 0° elevation was taken as a 0dB reference. ΔDI_{system} indicates how different the DIs of these systems are to the DI of the real sources and is given by

$$\Delta DI_{system} \leq \frac{1}{F} \sum_{1 \leq f \leq F} |DI_{system}(f) - DI_{loudspeakers}(f)| \qquad (2)$$

where F is the number of frequency bins, $DI_{system}(f)$ is the measurement of the DI for the 3D reproduction technique $system$ at frequency bin f and $DI_{realsources}(f)$ is the measurement of the DI for the reference $realsources$ at frequency bin f.

$\Delta DI_{system}(f)$ measured for the three manikin positions and the 8 systems are summarized in Table 1.

Description	Distance to the sweet spot		
	0cm	10cm	20cm
DBAP	1.61	1.15	**1.69**
HOA, PI, basic	2.50	2.50	2.99
HOA, PI, max_{RE}	2.27	3.08	3.33
HOA, EP, basic	2.02	2.55	2.72
HOA, EP, max_{RE}	1.89	2.96	3.82
VBAP	**0.96**	**0.92**	1.84

Table 1: Summary of the difference between the DIs measured for DBAP, HOA and VBAP systems and that measured with real sources. Values are in dB.

At the center

Figure 2 shows the DI for all the systems tested, measured at the sweet spot using a KEMAR manikin and hearing aid satellites. The DI obtained with VBAP at low frequencies is larger than that of the reference. This is caused by the integration of the DI over a cone of 60° around the 0° direction: when the DI is computed using only one direction versus all the others, the DI obtained with VBAP is always smaller than that obtained with the reference. A similar effect can be observed in Figure 2 for all HOA decoders; this is a consequence of a larger sound cancellation at the back with HOA at 250Hz and 320Hz. The increased DI is caused by the room effect, as truncating the impulse responses with a 237 samples window to remove room effect shows a lower DI at 250Hz and 320Hz, as shown in Figure 4. HOA sound reproduction produces a drop of DI above 2kHz, reaching its lowest point at 3150Hz. This could be caused by the errors of reconstruction: as explained above, at the position of the hearing aid (8.5cm off-centered on a KEMAR manikin), the sound field reconstruction is correct up to 1909 Hz.

Figure 3 compares the 2D polar pattern of the beamformer obtained for sound reproduction with real sources, VBAP, and HOA, at 500Hz and 3150Hz (the frequency at which the minimum DI was reached for HOA). The beamformer produces less directive patterns at high frequencies when used with HOA. A possible explanation could be the resulting delay between the signals. For comparison of

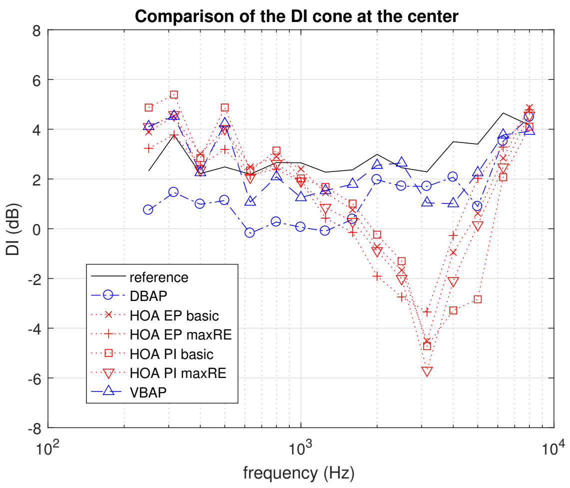

Figure 2: DIs measured with DBAP, HOA, VBAP and with real sources for a KEMAR manikin centered on the sweet spot.

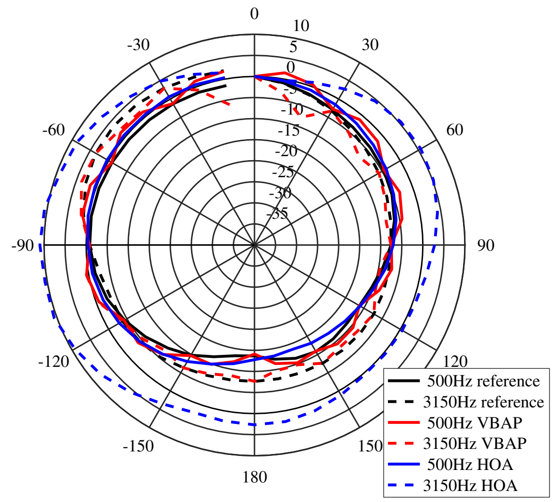

Figure 3: Comparison of polar plots of the beamformer, measured for real sources, for VBAP, and for HOA sounds. The polar pattern of the beamformer for real sources was interpolated between loudspeakers for this figure. Gains are in dB.

this delay, the delay between the signals was estimated at 500Hz and 3150Hz as a mean of the phase of the transfer function from the front microphone to the rear microphone over a bandwidth of 1/3rd octave. It was computed for the loudspeaker measurement, the VBAP measurement, and the HOA Pseudo-Inverse with basic weighting measurement. The delay was then estimated using the lag of the cross-correlation of the envelope of the impulse response for a source coming from 0° azimuth and 180° azimuth. At 500Hz, the effective delay between the signals is 31 μs / -39 μs (frontal / rear source) for real sources, 30 μs / -37 μs for VBAP, and 33 μs / -40 μs for HOA. Delays between the signals for the front and rear cases are close for the three types of sources compared to what happens at 3150Hz. The delay should be close to opposite when coming from the front and from the rear. A possible explanation of the difference is the non symmetrical shape of the head. At 3150Hz, the delay between the signals is 44 μs / -27 μs for real sources,

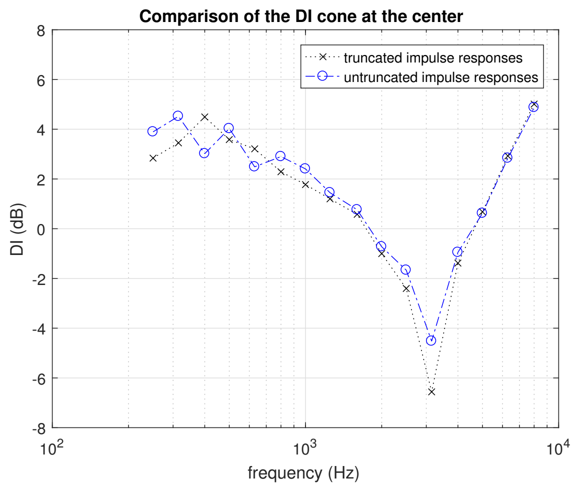

Figure 4: Comparison between the DI obtained with HOA (energy preserving, basic weighting) for truncated and untruncated impulse responses.

43 μs / -27 μs for VBAP, and -11 μs / -2 μs for HOA. For HOA, the measured delay is widely asymmetrical, too small compared to the real sources, meaning the attenuation of the beamformer will be insufficient. The absolute difference of DI between DBAP and the real sources is only marginally larger than between VBAP and the real sources. However, at low frequencies, DBAP is less directive than the real sources whereas both VBAP and HOA are more directive than the real sources.

Table 1 confirms these findings: the difference between the reference and VBAP is smaller than between the reference and various HOA decodings or between the reference and DBAP. Among the HOA decoders, pseudo-inverse decoding gives worse results than energy-preserving decoding. It is hypothesized that the sub-optimal positions of the loudspeakers could explain this difference. For comparison, the difference between the beamformer of the reference and a simulated ideal omnidirectional microphone at the center of the loudspeaker array would be $\Delta DI_{omnidirectional} = 2.89$ dB.

10cm off-centered

Figure 5 shows the DI for all the systems tested, measured with the KEMAR head 10cm off to the left of the sweet spot.

The HOA reconstruction error increases and the drop frequency of the beamformer's DI decreases. The minimum of the DI is reached at 2500Hz instead of 3150Hz when KEMAR is located at the center of the array. With all the HOA decoders, the DI also drops at high frequencies. Similarly to what was observed at the center, the energy-preserving decoding gives better results than the pseudo-inverse decoding. DBAP and VBAP give comparable results (ΔDI_{VBAP} =0.92 dB and ΔDI_{DBAP} =1.15 dB).

20cm off-centered

Figure 6 shows the DI for all the systems tested, measured with the KEMAR head 20cm off to the left of the sweet spot.

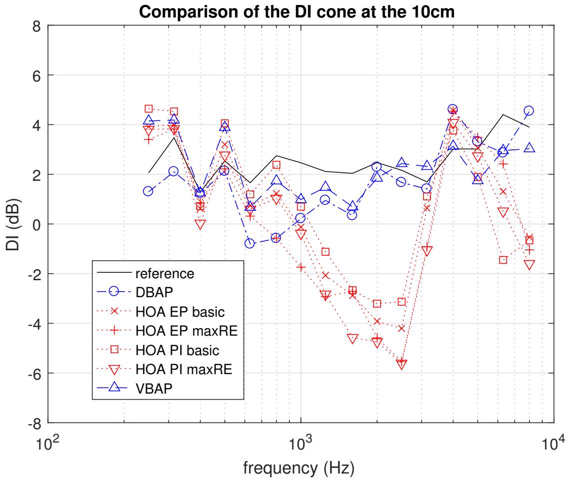

Figure 5: DIs measured with DBAP, HOA, VBAP and with real sources for a KEMAR manikin 10cm to the left of the sweet spot.

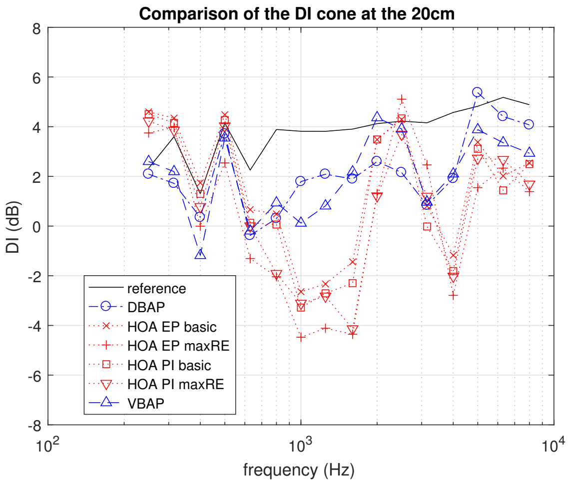

Figure 6: DIs measured with DBAP, HOA, VBAP and with real sources for a KEMAR manikin 20cm to the left of the sweet spot.

The HOA DI drop frequency decreases to 1kHz, with a second drop at 4kHz. The difference between the DI obtained with the 3D audio reproduction systems and the loudspeaker reproduction increases futher, and at 20cm the DI obtained with DBAP is the closest to the DI obtained with the loudspeakers.

Discussion

At the center, at low frequencies, the beamformer polar patterns obtained both with the VBAP and with the HOA systems were close to the real sources. Above 1000Hz, however, the DI of the beamformer when using HOA virtual sources drops. A similar effect could be observed in [11]. In Oreinos work, increasing the HOA order increased the limit frequency. This could not be observed here, the loudspeaker configuration being insufficiently regularly spaced for a correct fourth or fifth order HOA pseudo-inverse decoding. In [18], Oreinos concludes that the use of HOA imposes the use of a low-pass filter on the processed hearing aid signals. This is confirmed by the current study, and the theoretical cut-off frequency is influenced by the effect of the room in which the loudspeaker system is located and by the allowed movements of the listener, constraining the reliable evaluation of directional hearing aids using HOA. However, in this study, the hearing aid was always located at least 8.5cm from the center, the radius of KEMAR's head. For monaural objective measurements, one could imagine shifting the KEMAR manikin by 8.5cm, thus placing the ear and the hearing aid at the center of the system, which would increase the usable bandwidth.

In [6], Grimm used a beam pattern error as a measure of how suitable VBAP and HOA are for hearing devices evaluation. However, Grimm only studied 2D reproduction. In his study, using 12 horizontal loudspeakers (the 2D-equivalent of the 3D system used in the current study) would lead to a beamformer pattern error sufficiently low until below 2 kHz. Grimm's study was conducted simulating a sound reproduction system in an anechoic room, which explains the larger limit for the HOA sound reproduction, similar to the limit found by Oreinos. However, Grimm found that VBAP and HOA behave similarly, which is not the case in the current study. For comparison, Table 2 shows the criterion used by Grimm applied to the measurements of the current study.

Description	Distance to the sweet spot		
	0cm	10cm	20cm
DBAP	250	X	250
HOA, PI, basic	X	X	250
HOA, PI, max_{RE}	3150	1250	1000
HOA, EP, basic	250	X	250
HOA, EP, max_{RE}	X	1250	1000
VBAP	X	X	4000

Table 2: Frequency at which Grimm's error criterion is not verified anymore (in Hz). For these results, the beam patterns obtained with real sources had to be interpolated in order to compare it with each of the virtual source positions. An X means that the beam pattern error measurements for this condition were below 5.7dB at all frequencies from 250Hz to 6300Hz.

He found that for central, fixed head reproduction and a bandwidth of 4kHz, 18 loudspeakers are sufficient (8th order) for a reproduction on the horizontal plane, and that for off-centered position, 36 loudspeakers are necessary (17th order).

The results shown in this paper suggest a strong advantage for using VBAP or DBAP when evaluating hearing aids with beamformers. DBAP has been designed in the context of 3D audio reproduction for large audiences, where many listeners are not at the sweet spot. A consequence of this is the good performance of DBAP for the 20cm off-centered position. As discussed in the Introduction, VBAP causes sources to jump from loudspeaker to the other. In [19], Frank showed that Multiple-Direction Amplitude Panning (MDAP) provides a more linear virtual source position than VBAP, while maintaining a low number of active loudspeakers for any given source position. This should therefore be tested in the future.

The limited performance of HOA, particularly for off-centered positions, could be caused by the low HOA order and by a non-ideal loudspeaker positioning. For that reason, a future study will reproduce this experiment in a different room using a 32-loudspeaker setup as recommended by [27], and third and fifth order decoding will be compared.

In this study, no room simulation was added to the virtual sources. Similarly to the difference between a loudspeaker arrangement in an anechoic room and the same loudspeaker arrangement in a non-anechoic room, adding reverberation to the HOA and VBAP systems are expected to attenuate the differences between the systems. More particularly, without dereverberation engine, the efficiency of the beamformer will be decreased. The comparison between the three 3D audio reproduction methods with room simularion will be challenging: with HOA, it is possible to use room simulation softwares with HOA outputs [28, 29], HOA impulse responses recorded in situ or even first order Ambisonics impulse responses upmixed to HOA with Spatial Decomposition Method [30]. Solutions for VBAP and DBAP are more limited [31, 32].

More generally, out of the 5 types of audio processing included in hearing aids [12], only one was used here for the comparison of the 3D audio reproduction systems. A more systematic use of the various families of audio processing should be used for a better comprehension of the limitations of 3D audio for the evaluation of hearing aids.

In the current paper, differences were measured between the tested 3D audio reproduction systems, but the effect of these differences on perception remains unknown. Perceptual comparison between virtual sources and real sources with hearing aids was conducted using room simulation in [28], but was limited to HOA. Future work should therefore conduct similar test while including VBAP, MDAP and DBAP reproduction systems.

Conclusion

In this study, the behaviour of a hearing aid MVDR beamformer was compared using real sources, DBAP, HOA, and VBAP. It was shown that the difference between VBAP sources and real sources was smaller than between HOA sources and real sources at all three tested listening positions. VBAP is also a better solution than DBAP at the center and at the 10cm off-centered position. The findings of this study are similar to those of Oreinos and Grimm for HOA reproduction in anechoic rooms, but differ from Grimm's finding for VBAP. Future work should use a more regular loudspeaker array and add higher order HOA to the comparisons, complete the metrics with new metrics that could be used for characterising the usability of other hearing aid processings, and compare the results to those of perceptual evaluation tests.

Acknowledgments

This work was funded by CTI grant Nr. 25276.2 PFLS-LS. The authors would also like to thank Armin Stierli for his preliminary work on the 3D audio reproduction system used in this study.

References

[1] S. Kochkin. MarkeTrak VIII: Consumer satisfaction with hearing aids is slowly increasing:. *The Hearing Journal*, 63(1):19–20, January 2010.

[2] K. Payton. Using speech intelligibility metrics to predict the performance of hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 137(4):2271–2271, April 2015.

[3] K. C. Wagener and T. Brand. Sentence intelligibility in noise for listeners with normal hearing and hearing impairment: influence of measurement procedure and masking parameters. *International Journal of Audiology*, 44(3):144–156, March 2005.

[4] C. Oreinos and J. M. Buchholz. Objective analysis of higher-order Ambisonics sound-field reproduction for hearing aid applications. *Proceedings of Meetings on Acoustics*, 19(1):055023, June 2013.

[5] P. Minnaar, S. Favrot, and J. M. Buchholz. Improving hearing aids through listening tests in a virtual sound environment. *The Hearing Journal*, 63(10):40, October 2010.

[6] G. Grimm, S. Ewert, and V. Hohmann. Evaluation of Spatial Audio Reproduction Schemes for Application in Hearing Aid Research. *Acta Acustica united with Acustica*, 101(4):842–854, July 2015.

[7] P. Minnaar, S. F. Albeck, C. S. Simonsen, B. Sondersted, S. A. D. Oakley, and J. Bennedbaek. Reproducing Real-Life Listening Situations in the Laboratory for Testing Hearing Aids. Audio Engineering Society, October 2013.

[8] M. F. Mueller, A. Kegel, S. M. Schimmel, N. Dillier, and M. Hofbauer. Localization of virtual sound sources with bilateral hearing aids in realistic acoustical scenes. *The Journal of the Acoustical Society of America*, 131(6):4732–4742, June 2012.

[9] Jens Blauert. *Spatial Hearing - The Psychophysics of Human Sound Localization.* The MIT Press, rev sub edition, 1996.

[10] W. O. Brimijoin, A. W. Boyd, and M. A. Akeroyd. The Contribution of Head Movement to the Externalization and Internalization of Sounds. *PLoS ONE*, 8(12):e83068, 2013.

[11] C. Oreinos. *Virtual Acoustic Environments for the Evaluation of Hearing Devices.* PhD thesis, Macquarie University, Sydney, Australia, 2015.

[12] V. Hamacher, J. Chalupper, J. Eggers, E. Fischer, U. Kornagel, H. Puder, and U. Rass. Signal Processing in High-End Hearing Aids: State of the Art, Challenges, and Future Trends. *EURASIP Journal on Advances in Signal Processing*, 2005(18):152674, December 2005.

[13] G. Marentakis, F. Zotter, and M. Frank. Vector-Base and Ambisonic Amplitude Panning: A Comparison Using Pop, Classical, and Contemporary Spatial Music. *Acta Acustica united with Acustica*, 100(5):945–955, September 2014.

[14] S. Bertet, J. Daniel, E. Parizet, L. Gros, and O. Warusfel. Investigation of the Percieved Spatial Resolution of Higher Order Ambisonics Sound Fields: A Subjective Evaluation Involving Virtual and Real 3d Microphones. In *30th Audio Engineering Society International Conference: Intelligent Audio Environments*, March 2007.

[15] E. Corteel, L. Rohr, X. Falourd, K. V. NGuyen, and H. Lissek. Practical 3 dimensional sound reproduction using Wave Field Synthesis, theory and perceptual validation. In *Proceedings of the 11th French Congress of Acoustics and 2012 Annual IOA Meeting*, pages 895–900, Nantes, France, April 2012.

[16] S. Spors, H. Wierstorf, A. Raake, F. Melchior, M. Frank, and F. Zotter. Spatial Sound With Loudspeakers and Its Perception: A Review of the Current State. *Proceedings of the IEEE*, 101(9):1920–1938, September 2013.

[17] J. Cubick and T. Dau. Validation of a Virtual Sound Environment System for Testing Hearing Aids. *Acustica United with Acta Acustica*, 102:547–557, 2016.

[18] C. Oreinos, J. M. Buchholz, and J. Mejia. Effect of higher-order ambisonics on evaluating beamformer benefit in realistic acoustic environments. In *2013 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 1–4, October 2013.

[19] M. Frank. Localization Using Different Amplitude-Panning Methods in the Frontal Horizontal Plane. 2014.

[20] D. Kostadinov, J. D. Reiss, and V. Mladenov. Evaluation of distance based amplitude panning for spatial audio. In *In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal (ICASSP*, 2010.

[21] Ircam Spat URL: `http://forumnet.ircam.fr/product/spat-en/`, June 2017.

[22] M. Poletti, F. M. Fazi, and P. A. Nelson. Sound-field reproduction systems using fixed-directivity loudspeakers. *The Journal of the Acoustical Society of America*, 127(6):3590–3601, June 2010.

[23] F. Zotter, M. Frank, and H. Pomberger. Comparison of energy-preserving and all-round Ambisonic decoders. In *Fortschritte der Akustik, AIA-DAGA*, January 2013.

[24] J. Daniel. *Representation de champs acoustiques, applications Ã la transmission et Ã la reproduction de scenes sonores complexes dans un contexte multimedia.* PhD thesis, Universite Pierre et Marie Curie, 2001.

[25] D. B. Ward and T. D. Abhayapala. Reproduction of a plane-wave sound field using an array of loudspeakers. *IEEE Transactions on Speech and Audio Processing*, 9(6):697–707, September 2001.

[26] S. Moreau, S. Bertet, and J. Daniel. 3d Sound Field Recording with Higher Order Ambisonics - Objective Measurements and Validation of Spherical Microphone. In *120th Audio Engineering Society Convention*, Paris, 2006. Preprint 6857.

[27] Franz Zotter, Hannes Pomberger, and Markus Noisternig. Ambisonic Decoding with and without Mode-Matching: A Case Study Using the Hemisphere. In *2nd Int. Symposium on Ambisonics and Spherical Acoustics*, pages –, Paris, France, May 2010.

[28] C. Oreinos and J. Buchholz. Validation of realistic acoustic environments for listening tests using directional hearing aids. In *2014 14th International Workshop on Acoustic Signal Enhancement (IWAENC)*, pages 188–192, September 2014.

[29] S. Favrot and J. Buchholz. Validation of a Loudspeaker-Based Room Auralization System Using Speech Intelligibility Measures. Audio Engineering Society, May 2009.

[30] S. Tervo, J. Pätynen, A. Kuusinen, and T. Lokki. Spatial decomposition method for room impulse responses. *Journal of the Audio Engineering Society*, 61(1/2):17–28, 2013.

[31] Juha Merimaa and Ville Pulkki. Spatial impulse response rendering I: Analysis and synthesis. *Journal of the Audio Engineering Society*, 53(12):1115–1127, 2005.

[32] Ville Pulkki and Juha Merimaa. Spatial impulse response rendering II: Reproduction of diffuse sound and listening tests. *Journal of the Audio Engineering Society*, 54(1/2):3–20, 2006.